

Received October 9, 2020, accepted October 19, 2020, date of publication October 27, 2020, date of current version November 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3034144

Africa 3: A Continental Network Model to Enable the African Fourth Industrial Revolution

OLASUPO O. AJAYI¹, ANTOINE B. BAGULA¹, (Member, IEEE),
AND HLONIPHANI C. MALULEKE¹

Department of Computer Science, University of the Western Cape, Cape Town 7535, South Africa

Corresponding author: Olasupo O. Ajayi (olasupoajayi@gmail.com)

ABSTRACT It is widely recognised that collaboration can help fast-track the development of countries in Africa. Leveraging on the fourth industrial revolution, Africa can achieve accelerated development in health care services, educational systems and socio-economic infrastructures. While a number of conceptual frameworks have been proposed for the African continent, many have discounted the Cloud infrastructure used for data storage and processing as well as the underlying network infrastructure upon which such frameworks would be built. This work therefore presents a continental network model for interconnecting nations in Africa through its data centres. The proposed model is based on a multilayer network engineering approach, which first groups African countries into clusters of data centers using a hybrid combination of clustering techniques; then utilizes Ant Colony Optimisation with Stench Pheromone, that is modified to support variable evaporation rates, to find ideal network path(s) across the clusters and the continent as a whole. The proposed model takes into consideration the geo-spatial location, population sizes, data centre counts and intercontinental submarine cable landings of each African country, when clustering and routing. For bench-marking purposes, the path selection algorithm was tested on both the obtained clusters and African Union's regional clusters.

INDEX TERMS Africa, ant colony optimization, clustering algorithms, communication networks, computer networks.

I. INTRODUCTION

Despite the progress made toward actualizing the Sustainable Development Goals (SDGs), Africa still lags behind most of the world when it comes to socio-economic and industrial development. According to the United Nations' report on global human development index, as of today, only seven (7) of Africa's 55 countries have a Human Development Index (HDI) score greater than 0.75 [38]. The HDI is a metric used to measure the state of development of countries around the world. It takes into consideration factors such as level of healthcare, education, average life expectancy, among other things. It can clearly be inferred from this report that 75% of African countries are still developing, with poor infrastructure for health, education, transportation, electricity etc. Collaboration has been identified as a potential medium for accelerating development in many of these countries. A number of such collaborative efforts for

development of various aspects of Africa include: regional economic communities, such as the Economic Community of West African States (ECOWAS) and Southern African Development Community (SADC), for economic and political developments [7]; research and education networks such as the West and Central African Research and Educational Network (WACREN) and UbuntuNet, for education [39] and communication networks such as the Eastern Africa Submarine Cable System (ESSAy) and Melting Pot Indian-oceanic Submarine System (METISS) [19]. Besides these efforts, researchers have also proposed numerous frameworks for collaboration and resource sharing in Africa, such as those in the works of [1], [40], [41], [46]. Despite these proposals, only a few have focused on the underlying network architecture required to support such inter-continental collaborations. In [42], the authors considered an optical network to link Africa, while in [3], a dual ring network topology was proposed for interconnecting African countries. In [44], the role of the Internet on the evolution of communication infrastructure in 8 west African countries was considered.

The associate editor coordinating the review of this manuscript and approving it for publication was Halil Yetgin¹.

In [45], the author presented a holistic view of community networks in African countries, including the current state and various implementation barriers.

With respect to Clustering Algorithms (CAs), the authors in [13] surveyed the state of the art, discussed common distance metrics, as well as the similarities and evaluation functions of a number of CAs. They also presented different classes of traditional CAs, including: partition, hierarchical, fuzzy, distribution, density and graph; as well as modern CAs, such as kernel based, swarm intelligence, affinity and various hybrids. In [14], Cheng, *et al.* proposed a CA that is robust to noise. The algorithm was built upon a hierarchical model, with neighbour density used to filter out noise, while modified graph algorithms were used for clustering. In [15], an improvement on classic density-based CA was proposed to tackle the well-known issue of determining ideal scanning radius. The proposed model called DHeat incorporated an adaptive scanning radius mechanism. A number of other contemporary CAs exists but most are built upon the classic CAs, such as those in [16], [17]. In this work, we would stick with the classic CAs and apply them to clustering the African continent.

It is widely recognised that collaboration and digitisation can help achieve the SDGs by building upon the technologies of the fourth industrial revolution (4IR) to get a connected continent with potential for collaboration in health care services, industry and education. It can also deliver an Africa where infrastructure are shares in a federated way to bridge the gap between the poorest and the richest countries on the continent. A Federated Cloud model for health care in Africa was proposed in [1] as an attempt to achieve such goal. The work in [1] investigated both cooperative and competitive Cloud models and tested five different workload allocation schemes. Obtained results showed that cooperation could lower allocation delays and improve resources utilisation, while completion times and quality of service were better provided using the competitive model. Despite proposing a potential model, no discussion was made on the implementation of such a model, especially from a network architecture perspective. Prior to the work done in [1], Huang, *et al.* had investigated the viability of a network, referred to as “Africa 2”, that covers the entire African continent [3]. They proposed a dual-ring network architecture to service the 55 countries (54 plus Western Sahara*) of Africa [7]. In their work, they selected some countries as “bridge-nodes” with links between these nodes serving as connection points between the two rings. The work also compared a greedy algorithm and simulated annealing algorithm for selecting the optimal placement of the bridge node. Furthermore, they considered various models for achieving inter-connectivity among African countries, while factoring in growth potentials.

This article revisits the issue of African connectivity and proposes a continental network model called “Africa 3”, that can enable the African fourth industrial revolution through digitisation. In this work, we seek to address similar routing

objectives as “Africa 2”; however, by applying a completely different hierarchical approach. In our model, neighbouring countries are grouped into clusters forming digital islands. These clusters are then interconnected by a network engineered model based on a modified ant-colony optimisation algorithm. Africa 3, also borrows the federated approach from [1], which enables efficient connection of Cloud infrastructures within the continent.

The specific contributions of this work are:

- 1) Development of a continental network model for inter-communication across Africa. This model takes into consideration the technological capabilities (in terms of Cloud data centres), the population and geo-political importance of each African country.
- 2) Division of African countries into functional clusters using a number of clustering algorithms. This enables easy “visualisation” of each feature’s potential influence on the continent.
- 3) The proposal of a modified Ant Colony Routing algorithm for the identification of the best network routes within and between clusters of countries across Africa. The algorithm is based a combination of dynamic stench pheromones and weighted cost function.

With respect to the organization of this article, following this section is section II, with a review of related literature and the introduction of our proposed continental network model for Africa. In section III, we consider various clustering algorithms for the lower layer of Africa 3, while network routing at the upper layer is presented in section IV. Simulations and discussions of experimental results are done in section V, while open implementation challenges are highlighted in section VI. We then conclude the paper and give insights into future research directions in section VII.

II. THE AFRICAN NETWORK MODEL

As suggested earlier, “Africa 3” is an continental network that combines the routing objectives of “Africa 2” [3] with the Cloud federation approach proposed in [1]. Africa 3 consists of a hierarchical network, wherein at the lower level, neighbouring countries are grouped into clusters and the cluster heads or regional hubs are selected as the “bridge-node”. These bridge-nodes are then linked together, forming a continent-wide network at the higher level. The choice of which country is selected as a cluster head is based on population size, number of data centres, geographical location and regional importance. Table 1 puts these in perspective and shows the number of data centres in African countries [2], as well as the approximate population size and centroids of each country [4]. On Table 1, data centres values of zero implies that the country either does not have any data centre or the ones present therein are below the standard stipulated in [6].

Finding the optimal or “best” path in a network environment has been found to be an NP-hard problem in many cases and one which cannot be solved in intrinsic computing time. A number of meta-heuristic approaches have been developed

TABLE 1. Geo-demography of African countries.

SN	Country	Sub-Region	Centroid Coordinates	Pop. Est.	DC Count
1	Burundi	Central Africa	-3.35939665747, 29.8751215645	9,823,828	0
2	Cameroon	Central Africa	5.69109848986, 12.7396415575	21,917,602	0
3	Central African Republic	Central Africa	6.56823297048, 20.4682683098	3,859,139	0
4	Chad	Central Africa	15.333337585, 18.6449251291	11,039,873	0
5	Democratic Republic of Congo	Central Africa	-2.8774628897, 23.643961066	86,026,000	1
6	Equatorial Guinea	Central Africa	1.7055135464, 10.3413792389	1,222,442	0
7	Gabonese Republic	Central Africa	-0.586600249551, 11.7886286968	1,802,278	0
8	Republic of the Congo	Central Africa	-0.837874631184, 15.2196576195	3,697,490	0
9	Sao Tome and Principe	Central Africa	0.443914449308, 6.72429657927	179,200	0
10	Comoros	Eastern Africa	-11.8778344355, 43.6825396848	806,200	0
11	Djibouti	Eastern Africa	11.7487180559, 42.5606754026	864,618	0
12	Eritrea	Eastern Africa	15.361866184, 38.8461701125	6,536,000	0
13	Ethiopia	Eastern Africa	8.6227867931, 39.6008009763	112,078,730	0
14	Kenya	Eastern Africa	0.599880215445, 37.7959397293	45,533,000	6
15	Madagascar	Eastern Africa	-19.3718958677, 46.7047367428	22,434,363	1
16	Mauritius	Eastern Africa	-20.2776870433, 57.5712055061	1,261,208	9
17	Rwanda	Eastern Africa	-1.99033831693, 29.9198851524	10,515,973	0
18	Seychelles	Eastern Africa	-4.66099093522, 55.4760327912	90,945	0
19	Somalia	Eastern Africa	4.75062876, 45.70714487	12,316,895	0
20	South Sudan	Eastern Africa	7.30877944922, 30.2479000186	8,260,490	0
21	Tanzania	Eastern Africa	-6.27565408332, 34.8130998093	51,046,000	1
22	Uganda	Eastern Africa	1.27469298731, 32.3690797137	34,856,813	0
23	Algeria	Northern Africa	28.1589384945, 2.6173230092	40,100,000	1
24	Egypt	Northern Africa	26.4959331064, 29.8619009908	98,002,045	11
25	Libya	Northern Africa	27.0309449491, 18.0086616872	5,298,152	0
26	Morocco	Northern Africa	29.8376295469, -8.45615794798	33,337,529	5
27	Sudan	Northern Africa	15.9903566856, 29.9404681199	40,235,000	0
28	Tunisia	Northern Africa	34.1195624649, 9.55288358933	10,982,754	2
29	Western Sahara*	Northern Africa	24.2295673865, -12.2198275485	510,713	0
30	Mauritania	Northern Africa	20.257367056, -10.3477981455	3,718,678	0
31	Malawi	Southern Africa	-13.2180808806, 34.2893559871	16,832,900	0
32	Mozambique	Southern Africa	-17.2738164259, 35.5336754259	28,013,000	0
33	Zambia	Southern Africa	-13.4582415186, 27.7747594637	15,473,905	0
34	Zimbabwe	Southern Africa	-19.0042041882, 29.8514412019	13,061,239	1
35	Angola	Southern Africa	-12.2933605438, 17.5373676815	24,383,301	3
36	Botswana	Southern Africa	-22.1840321328, 23.7985336773	2,024,904	0
37	Eswatini	Southern Africa	-26.5584304502, 31.4819369011	1,119,375	0
38	Lesotho	Southern Africa	-29.5800318833, 28.2272313098	1,894,194	0
39	Namibia	Southern Africa	-22.1303256842, 17.209635667	2,280,700	0
40	South Africa	Southern Africa	-29.0003409534, 25.0839009251	58,558,270	21
41	Benin	Western Africa	9.64175970225, 2.32785254449	10,008,749	0
42	Burkina Faso	Western Africa	12.2695384574, -1.75456600914	18,450,494	0
43	Cabo Verde	Western Africa	15.9552332421, -23.9598882026	491,875	0
44	Gambia	Western Africa	13.4496524352, -15.3960129463	1,882,450	0
45	Ghana	Western Africa	7.95345643541, -1.21676565807	27,043,093	2
46	Guinea	Western Africa	10.4362159331, -10.9406661161	10,628,972	0
47	Guinea Bissau	Western Africa	12.0474494815, -14.9497244459	1,530,673	0
48	Cote d'Ivoire	Western Africa	7.62842620235, -5.5692156998	22,671,331	0
49	Liberia	Western Africa	6.45278491657, -9.3220757269	3,476,608	0
50	Mali	Western Africa	17.3458158135, -3.54269064851	14,528,662	0
51	Niger	Western Africa	17.4191249296, 9.38545881539	17,138,707	0
52	Nigeria	Western Africa	9.59411452233, 8.08943894771	200,963,599	10
53	Senegal	Western Africa	14.3662417313, -14.4734923973	14,354,690	0
54	Sierra Leone	Western Africa	8.56329593038, -11.7927124668	6,348,350	0
55	Togo	Western Africa	8.52531356141, 0.962328449769	6,191,155	0

to solve the optimal path finding problem, prominent among which is the Ant Colony Optimization (ACO). Ant Colony Optimization (ACO) is a meta-heuristic algorithm that draws inspiration from the self-organizing behaviour of ants [20]. Like in real ants which randomly search for food and lay pheromones to guide other ants to food sources, ‘ants’ in ACO also update pheromone levels along traversed pathways. These artificial ants have memories, where they keep information about previously visited paths and local solution. Upon returning to the ‘nest’ from food foraging, each ant updates the pheromone level on the trail it traversed. These pheromone levels are used to guide other ants in subsequent

iterations. ACO has been modified to solve a diverse class of NP-hard problems, such as finding shortest paths in packet switched networks [32]–[34] and finding best routes in vehicle routing problems, as done in [35] and [37].

The relatively limited number of research works directly focusing on network routing across Africa is an indication of a research gap and serves as motivation for this article. In this work, we propose a continental network for intercommunication across African nations and use ACO for optimal path selection across these nations. Such a network can serve as a backbone for numerous collaborative health and educational frameworks being proposed for Africa. Our proposed model,

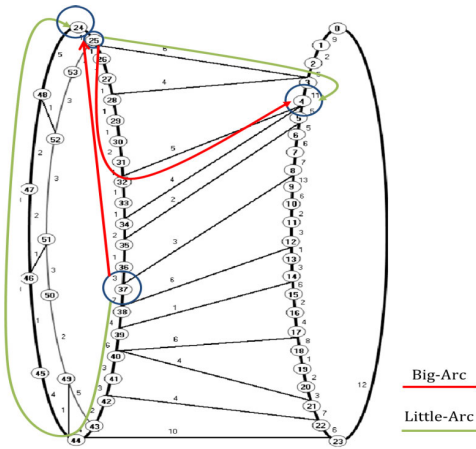


FIGURE 1. Two-Ring Network showing sample routing paths [3].



© 2017. Sahel and West Africa Club Secretariat (SWAC/OECD)

FIGURE 2. African countries grouped into regions [8].

called Africa 3, groups African countries into clusters, similar to that of the African Union (Fig. 2) and then applies ACO to identify the most efficient network path within and across clusters. In comparing our model to the dual-ring model proposed in [3], rather than connecting individual countries directly to the dual-ring, our model connects clusters of countries together. The benefits of this approach are as follows:

- **Avoidance of single point of failure:** with the ring topology, every node is potentially a point of failure. This is of particular concern in many African countries, where stable and uninterrupted electricity supply is still a challenge. This exponentially increases the probability of nodal failures. Though network traffic can flow in both directions in such dual-ring topology, it is not uncommon for multiple countries in Africa to experience total power failure.
- **Reduced routing distance:** in [3], the Little-Arc-First (LAF), shown in Fig. 1, was the primary routing path as it represented the shortest network distance. This might not always be the best option, as is the case of transmission between Dakar (node #25) and Abuja (node #4), where the LAF is the path: #25, #3, #4 and costs $6 + 11 = 17$. The Big-Arc on the other hand goes through the path: #25, #26, #27...#32, #4 and has a cost of only 13. Similarly, when routing from Luanda (#37) to Nouakchott (#24), the LAF is the path through #38, #39 ...#48, #24 at a cost of 43, while the Big-Arc through path #36, #35 ...#25, #24 only costs 17. Using the ACO, our model ensures that the path selected between countries is always the best (lowest route cost).

A. THE AFRICAN CONTINENTAL NETWORK MODEL (AFRICA 3)

In this subsection we present our model for interconnecting Africa. As stated earlier, it is a two level model with levels described as follows:

Lower Level We begin by grouping countries in Africa into clusters, using a combination of geographical coordinates (latitude and longitude), population size, number

of data centres and number of independent intercontinental submarine cable landings present in each country. These features were chosen because: i. geographical coordinates are one of the natural ways of clustering countries; ii. population size and socio-economic power influence the demand for network traffic; iii. data centre count gives a quick insight to the level of technological development, with the world rapidly adopting numerous technologies that are dependent on Cloud computing [43]; iv. submarine cable landings are connection points that link Africa to the rest of the world. These points therefore serve a vital role of gateways in our model.

Upper Level At this level, our model focuses on interconnecting the islands of clusters formed at the lower level. Upon successfully clustering the continent, we apply a variant of ACO called Ant-Colony Routing with Stench Pheromones (ACO-SP) to find optimal network paths within clusters and to gateways. Of particular importance is the fact that, while searching for optimal routes, our model strives to avoid paths that directly cross the Sahara desert.

B. ARCHITECTURAL FRAMEWORK

Fig. 3 gives an illustrative overview of our proposed continental network model for Africa. At the first phase, information about the 55 African countries are obtained. These information include: coordinates of each country’s centroid (geographical mid-point), population size and number of data centres present. They were obtained by web scraping and/or API calls on [2], [4], [5]. These information are then fed into different clustering models as features. Here the countries are grouped into clusters, with the Africa Union’s clustering used as a benchmark. This clustering phase is discussed in section III.

Upon successfully clustering the continent, the various clusters are fed into a routing model, where optimal paths interconnecting countries within each clusters (intra-cluster) and across clusters (inter-cluster) are obtained. The routing model is based on ACO meta-heuristic, and is discussed in section IV. The final output is the African Continental Network.

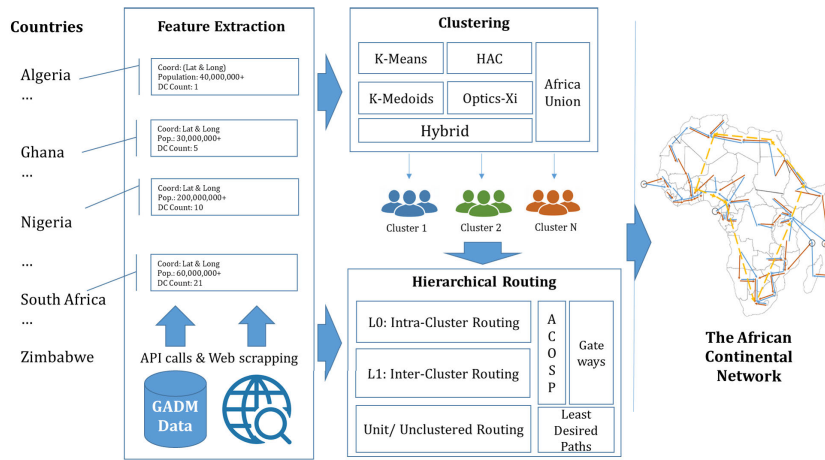


FIGURE 3. The African Continental Network Framework.

III. AFRICA IN CLUSTERS

In this section, we focus on the lower level of our Africa 3 model and consider various algorithms for effectively splitting African countries into clusters. The African Union has grouped the continent into 5 regions, namely – Central, Eastern, Northern, Southern and Western, with member countries listed on Table 1 and shown in Fig. 2. In grouping countries into clusters, we considered a number of unsupervised machine learning algorithms and use the AU’s clustering layout as a comparative reference.

Table 1 lists our input data and consists of two types of data, viz.: geo-spatial data (latitude and longitude) and temporal data (population size and data centre count). The final feature of inter-continental submarine landings would be discussed later. In clustering these data, we considered two types of K-means clustering algorithms, a hierarchical clustering algorithm and a density-based clustering algorithm. These algorithms are described as follows:

- **Classic K-means (Lloyd-K-Means):** The classic K-Means is a centroid based clustering algorithm, which, in essence, implies that it determines cluster members by how close data points are to a centre point [22]. For k-means, we initially set the value of k to 5, in line with the AU’s regional distribution of the continent into five regions. Rather than solely relying on the AU’s clustering for the value of k, we also used two elbow methods - the Distortion score and Calinski-Harabasz score [31] to determine the optimal value of k. We then repeated the clustering using this new optimal value of k. Euclidean distance [23] was used as proximity metric for selecting cluster members. Our decision to use the Euclidean distance was because Africa is uniquely located almost symmetrically around the equator (latitude 0°) and the prime meridian (longitude 0°), at 30°W, 60°E, 40°N, 40°S; the use of the Euclidean distance is therefore justifiable as there are no longitudinal wrap around (0° and 180°). Despite this

justification and a similar other [9], for thoroughness and avoidance of doubt, we also clustered using the Great Circle Distance (Haversine) [24]. We implemented this with the k-Medoids algorithm described next.

- **K-Medoids:** K-Medoids or Partitioning Around Medoids (PAM) algorithm, like K-means is also a partitioning algorithm which attempts to group similar data together based on distance. However, unlike k-means the cluster centre in K-Medoids is one of the data points rather than an average of the points in the cluster as is the case with k-means. K-Medoids also allows the use of arbitrary distance metric [21]. With the earth being spherical, it has been argued that the Euclidean distance might be less than ideal for clustering geo-spatial data. To this end, we implemented the K-Medoids algorithm with the Haversine distance metric [24], described in (1).

$$\begin{aligned}
 \text{distance} &= R * c \\
 c &= 2 * \text{atan2}(\sqrt{a}, \sqrt{1-a}) \\
 a &= \sin^2\left(\frac{\text{lat}_2 - \text{lat}_1}{2}\right) \\
 &\quad + \cos(\text{lat}_1) * \cos(\text{lat}_2) * \sin^2\left(\frac{\text{lon}_2 - \text{lon}_1}{2}\right)
 \end{aligned}
 \tag{1}$$

where lat_i and lon_i are the latitude and longitude of the respective countries i in Radians and R is the radius of the earth.

- **Hierarchical Agglomerative Clustering (HAC):** This is a hierarchical “bottom-up” clustering technique that treats each data point (country) as a unit cluster and pairs up similar clusters. This process is repeated until all nodes are grouped into the minimum number of clusters [11]. We used two pairing metrics – the Euclidean distance and Complete Affinity. We chose the Complete Affinity because we sought to maximize the distance between data points. The advantage of this becomes more apparent when we cluster based on the geo-spatial,

population and data centre sizes; as having multiple countries with large population sizes within the same cluster is undesirable. Similarly, for data centres, it is also desirable to have widely (and evenly) dispersed data centres across the continent, in order to improve the overall technological capabilities of the continent as a whole.

- **OPTICS-Xi Clustering:** The Ordering Points To Identify the Clustering Structure-Xi (OPTICS-Xi) algorithm [25] is based on the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) clustering algorithm, which has been reported to be better for clustering geo-spatial data. OPTICS-Xi is a variant that relaxes the need to set correct values of epsilon (ϵ), which is often difficult to determine for most data sets [25], [26]. For our implementation, we considered a country as a point P and then used the OPTICS-Xi clustering to find the set of countries that were within a “circular radius” (ϵ) of P . Unlike the DBSCAN, OPTICS-Xi relaxes the need to set a value for epsilon. However, we set minPts (neighbourhood density) to 3, thereby ensuring that each cluster had at least 3 countries.
- **Multi-Feature Clustering:** We combined all features from our dataset and cluster based on this. Two clustering techniques were used here: i. a hybrid of K-Medoids and HAC; ii. OPTICS-Xi. For this combined feature clustering we replace both Euclidean and Haversine distances with the weighted distance defined in (2).

$$\text{weightedDistance} = \alpha * \text{hdistance} + \beta * \text{population} + \gamma * \text{dsize} \quad (2)$$

where hdistance is the haversine distance between two countries, population is the population size and dsize is the number of data centres present. α, β, γ are weight factors, which respectively determine the influence of hdistance , population and dsize . Since the 3 features were of different scales, we normalized them using the min-max normalization [27], [28] defined in (3). Resulting values were scaled between 0 (relative min. value) and 1 (relative max. value).

$$X' = \frac{x - \min(x)}{\max(x) - \min(x)} \quad (3)$$

Once the clusters have been obtained, we then used ACO-SP for path finding. The processes involved are described in the next section.

IV. CONTINENTAL NETWORK GRID

One of the primary objectives of this work is to propose a network to interlink countries in Africa for the purpose of collaborating and sharing network infrastructure. Recent evidence has shown an imbalance in the level of development of Africa countries. Some countries are comparatively at the forefront of innovation and technological advancements, while others are stuck in the middle age. By leveraging on collaborative

technologies, growth and advancements in many developing countries of Africa can be accelerated. Such collaboration could enhance growth in various areas including but not limited to health care, trans-boundary water management and general socio-economic activities. However, in order to achieve this, an underlining and robust network needs to be put in place.

A. SELECTING CONTINENTAL GATEWAYS

We define continental gateways as landing points for trans-continental submarine fibre optics cables, which provide Internet connectivity to the continents. There are a number of continental gateways across Africa, most of which are located in coastal cities. Table 2 gives a concise summary of these continental gateways in Africa [19]. The table shows that there are about 13 cities with trans-continental submarine fibre optics cables terminals in Africa. Note that, though EASSy has been included, it is not a trans-continental fibre optic line as it does not link Africa to any other continent. These continental gateways would serve as end-point for our continental network grid. There are a number of them across Africa, among which we choose a few, which we call Prime Continental Gateways (PCGs). For this work, a PCG is defined as any African country having multiple independent submarine cables landing within her shores. Using Table 2 as a guide, we chose countries with 5 or more trans-continental submarine fibre optics cables landings. These include: Algeria (5), Nigeria (8), Cameroon (6), South Africa (8), Kenya (6), Djibouti (9), Egypt (15), and Libya (5). Though Cameroon and Nigeria are in close proximity, the large population (and by extension network demand) of Nigeria justifies the selection of Cameroon as a PCG.

B. TRAFFIC ROUTING

We apply a modified ACO to find optimal network paths across the continent. This ACO is applied twice in our continental model; first at the cluster level and then at the continental level. At the lower hierarchy (cluster level), ACO is used to find paths from cluster members to the cluster head within a cluster. At the continental level, it is used to find ideal paths between the various cluster heads and the PCGs. For comparative purposes, we also considered an ‘unclustered’ approach, wherein ACO is applied directly at the continental level. In this model, all 55 member countries of Africa (54 + Western Sahara), are taken as nodes and ACO is applied to obtain an optimal network paths across the entire continent and to the PCGs.

In developing this continent-wide network for Africa, we considered a variant of ACO called Ant Colony Optimization with Stench Pheromone (ACO-SP) [36]. ACO-SP has been used in network packet routing with the primary objective of preventing congestion. In a typical ACO algorithm, the ants obtain the best route (often the shortest route) across a network. The problem with this, in a packet routing network, is that all traffic would flow through this “best

TABLE 2. Major trans-continental submarine fibre optics landings in Africa.

S/N	Name	Source	Destination	Counties Traversed	Country Count
1	MainOne	Portugal	Nigeria (Lagos)	Ghana	3
2	SAT3 / WASC	Spain	South Africa (Western Cape)	Senegal, Côte d'Ivoire, Ghana, Benin, Nigeria, Cameroon, Gabon and Angola	12
3	South Africa Far East (SAFE)	South Africa (Western Cape)	Malaysia (Penang)	Africa (Mauritius), Réunion and India	5
4	West Africa Cable System (WACS)	United Kingdom	South Africa (Western Cape)	Africa (Namibia, Angola, DRC, Congo, Cameroon, Nigeria, Togo, Ghana, Côte d'Ivoire, Cape Verde, Canary Islands), Portugal	14
5	Seacom	South Africa (Cape Town)	India (Mumbai)	Africa (Tanzania, Djibouti, Kenya Mozambique, Egypt, Botswana, Uganda, Rwanda), Saudi Arabia, India, France, Germany, Netherland, UK.	16
6	EASSy	South Africa (KwaZulu-Natal)	Sudan	Tanzania, Djibouti, Mozambique, Somalia, Kenya, Comoros, Madagascar	9
7	The East African Marine System (TEAMS)	Kenya (Mombasa)	UAE (Fujairah)		2
8	South Atlantic Inter Link (SAIL)	Brazil (Fortaleza)	Cameroon (Kribi)		2
9	South Atlantic Cable System (SACS)	Brazil (Fortaleza)	Angola (Sangano)		2
10	Asia Africa Europe (AAE-1)	China (Cape D'Aguiar)	France (Marseille)	Vietnam, India, Cambodia, Oman, Thailand, UAE, Myanmar, Qatar, Malaysia, Italy, Pakistan, Yemen, Saudi Arabia, Greece, Africa (Egypt, Djibouti)	18
11	Europe India Gateway (EIG)	India (Mumbai)	UK (Bude)	Oman, UAE, Saudi Arabia, Gibraltar, Monaco, Portugal, Africa (Egypt, Libya, Djibouti)	12
12	SeaMeWe-3	Australia	Europe	Numerous Asian and European Countries + Africa (Egypt, Djibouti)	28
13	SeaMeWe-4, SeaMeWe-5+	France (Toulon)	Malaysia (Melaka)	European, Euro-Asian countries, Africa (Algeria+, Tunisia+, Egypt and Djibouti)	15, 17+
14	FALCON	Egypt (Suez)	India (Mumbai)	Iraq, Yemen, Saudi Arabia, Kuwait, Oman, Iran, Qatar, UAE, Bahrain, Africa (Sudan)	12
15	PEACE	Kenya (Mombasa)	France (Marseille)	Africa (Egypt, Somalia, Djibouti, Seychelles), Pakistan	7

route”, thereby: i. under-utilizing other routes, ii. resulting in network congestion on the optimal route. ACO-SP, therefore, introduces a second pheromone called “stench pheromone” used to discourage selection of a given path, thus pushing

TABLE 3. Parameters and settings.

Variable	Value	Description
f	0 – 1	Used to introduce randomness into the system
T	0.8	Threshold value
s	1 – 55	Current location of ant
d	1 – 55	Next location of ant
k	1 – 1000	Individual Ant
V	1 – 55	List of countries visited by ant k
Φ	0.2	Initial Pheromone level
α	0.33	Determines the influence of distance
β	0.33	Determines the influence of population
γ	0.33	Determines the influence of DC count
H_{sd}	Equation (1)	Haversine distance between two countries s and d
Q		Population of destination country
C		Data centre counts of destination country
Δ	Equation (2)	Route cost (weighted distance) between two countries.
ρ	0.2	Pheromone evaporation rate
P_{sd}		Probability of moving from country s to country d
G	[PCGs]	Set of desirable paths
U		Set of undesired paths

some ants away from the best path and forcing them to create alternate best path(s).

Applying this concept, we created a list of undesirable paths, which includes paths that directly traverse the Sahara desert. Conversely, we also created a list of desired paths, which are paths we encourage the ants to traverse and consists of the PCGs. In implementing this, we created a multiplicative index, which either increases the rate of pheromone evaporation for undesired paths or reduces it for desired paths.

Our implementation process is as shown in Algorithm 1 with parameters detailed on Table 3. On the table, f is used by the ACO-SP algorithm to introduce a level of randomness into the system. s and V are variables which indicate the location of ants; their values ranges between 1 and 55, with each location representing an African country. 1000 ants were used in the experiment and this is indicated with variable k . Δ is a variable representing the weighted distance between two countries and is defined by 2. ρ is a factor which determines how fast the pheromones evaporate. It is a unique feature that distinguishes our work from [36] and is defined in step 5, equation 8 of Algorithm 1. Finally, S and U are sets of countries with PCGs and within the Sahara desert respectively.

V. SIMULATIONS

Simulation were carried out on a system configured with Intel core i7 processor, 8GB RAM and Windows 10 Operating System. Python (Jupyter notebook with Scikit Learn library [30]) and Java (ELKI Data mining framework for Java [29] with Eclipse IDE) were used, as both offered complementary libraries and visualization tools.

Geo-spatial data of each country were scrapped from [5], while data on population were obtained from [4] and DC size from [2]. After normalizing (3), the data were passed into the respective clustering module. K-Means, HAC and Optic-Xi clustering were done in Python, while K-Medoids and the combined clustering were done using ELKI. This was because ELKI supports K-Medoids and also enables the use of custom distance function (such as 2), which was vital for this work. However, a major disadvantage of ELKI was

Algorithm 1 Modified Ant Colony Routing With Stench Pheromones

1. Each ant starts from a randomly selected country.
2. Generate a random number f , $f = [0, 1]$ and a threshold value (0.8).
3. From any given country (s), the ants choose the next country (d) to traverse to using (4) adapted from [37]. This strives to find the maximum ratio between pheromone level and distance as follows: If $f \leq threshold$

$$d^k = \max\left(\frac{\Phi_{sd}}{\Delta_{sd}}\right), \quad d^k \notin V^k \quad (4)$$

In calculating our route cost Δ , we used a weighted distance metric based on (5), where $\alpha + \beta + \gamma = 1$.

$$\Delta_{sd} = \alpha * H_{sd} + \beta * Q_d + \gamma * C_d \quad (5)$$

If $f > threshold$ the next country is calculated based on a probabilistic function which favours the shortest path with the highest pheromone level. This probability is given in (6) described as follows:

$$P_{sd} = \frac{\Phi_{sd}}{\Delta_{sd} * \sum_{su} \frac{\Phi_{su}}{\Delta_{su}}}, \quad u, d \notin V^k \quad (6)$$

4. Update Pheromone level based on (7) described as:

$$\Phi_{sd} = (1 - \rho_{sd}) * \Phi_{sd} + \frac{\rho_{sd}}{\Delta} \quad (7)$$

5. Adjust evaporation rate: in our case evaporation rate is not constant across the entire system but rather varies for specific trails to encourage or discourage the use of certain trails. Evaporation rate is thus adjusted based on the conditions given in (8).

$$\rho_{sd} = \begin{cases} \rho, & d \notin G, U \\ \rho * 0.25, & d \in G \\ \rho * 0.75, & d \in U \end{cases} \quad (8)$$

G and U are respectively sets of encouraged trails (such as PCGs) and set of discouraged trails (trails through deserts).

that its maps were in gray scale and it only gave cluster points. We therefore had to manually colour code each of the resulting clusters, and superimpose the outline of the map of Africa.

A. CLUSTERING

As described above, our Africa 3 model groups African countries into clusters. For clustering, we compared four different clustering algorithms namely K-Means, KMedoids, OPTICS-Xi and HAC; using two distance metrics - Euclidean distance and great circle distance (Haversine). We carried out our clustering in multiple phases. At the first phase, we clustered using only geometric values (i.e. Latitude and Longitude). Results of this phase are shown in Fig. 5 to 8. We then clustered based on DC size and populations. Being

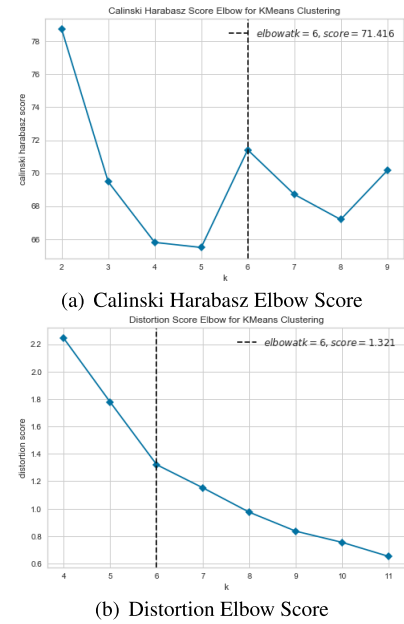


FIGURE 4. Selecting ideal value of K.

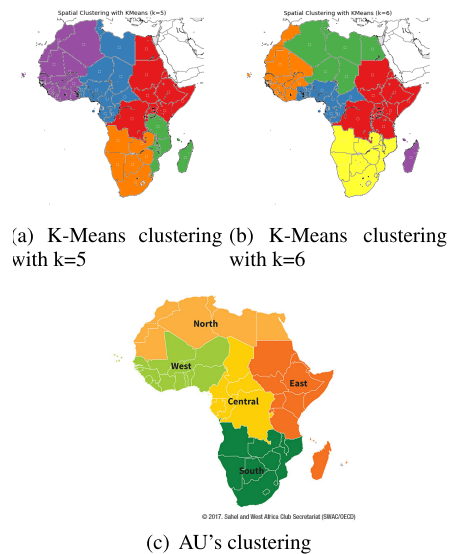


FIGURE 5. K-means clustering using Euclidean distance.

1 dimensional, the results were simple histograms of country versus frequency (count). These are not shown for purpose of space savings. Afterwards, the three features were combined for clustering and obtained results are shown in Fig. 9(a), 9(b) and 9(c). The AU clustering map, has been included in each result for reference purposes.

1) CLUSTERING USING K-MEANS

Fig. 5(a) shows clustering results when K-Means is used, with the value of k set to 5 and the Euclidean distance used. The choice of $k = 5$ was done purely for the purpose of comparing our results with AU's which also has 5 clusters. In Fig. 5(b), rather than fully relying on the AU's classification and setting k to 5, we instead chose the value for k by running two

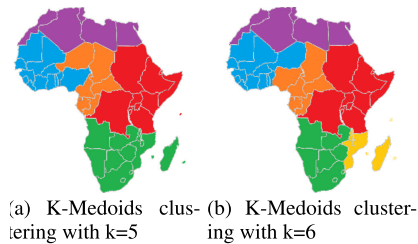


FIGURE 6. K-Medoids clustering using Haversine distance.

elbow methods - the Distortion score and Calinski-Harabasz score [31]. Both of them gave a value of 6 for k , as shown in Fig. 4(a) and 4(b). For both values of k (5 and 6), the resulting cluster maps differ from that of the AU's. This is most prominent in the Northern and Western clusters.

2) CLUSTERING USING K-MEDOIDS

We then proceeded to cluster based on the great circle distance, a metric which takes the shape of the earth into consideration, when measuring the distance between points. However, it has been reported that K-Means is not well suited for use with distance metric other than Euclidean, for this purpose we also clustered using K-Medoids. Obtained results are depicted in Fig. 6(a) and 6(b), for values of $k = 5$ and 6 respectively.

For both values of k , obtained results were similar to the AU's clustering. When k was set to 5, Fig. 6(a) shows Madagascar, Comoros and Mauritius being grouped in Southern Africa. The DRC was classified as part of Eastern Africa, while Niger was removed from Western Africa and included in Central Africa. Finally Mauritania was classified as a West African country.

With k set to 6, similar results were obtained for Northern and Eastern Africa, with the only exception being Seychelles which was put in a separate cluster, alongside Madagascar, Comoros, Mauritius, Mozambique and Malawi. To the West, Nigeria was classified as a Central African country, while Niger took its place in West Africa. These are as shown in Fig. 6(b).

3) CLUSTERING USING HAC

We then considered a third clustering algorithm - the HAC, which is an Agglomerative clustering technique. The advantages of this algorithm are its efficiency with small datasets and its implicit ability to determine the ideal number of clusters. Fig. 7 shows the Dendrogram, from which we chose distances of 35 and 50. This translated to 5 and 6 clusters, with the resulting cluster diagrams shown in Fig. 8. Results obtained in both cases were highly skewed towards Central Africa, with Northern and Eastern Africa having only 4 countries each.

4) MULTI-FEATURE CLUSTERING

From the above results, only the K-Medoids algorithm gave results closest to the AU's clusters; for this purpose we

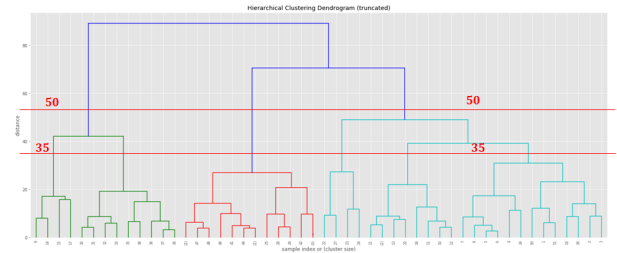


FIGURE 7. HAC Dendrogram (Distance = 35, 50).

decided to stick with it. Though we also clustered based on population and data centre size, the obtained histograms are not shown. We instead show the result of combining all three features (geographical coordinates, population and data centre size).

We compared the K-Medoids algorithm (for k values of 5 and 6) based on the combined featured and Haversine distance metric. Figs. 9(a), 9(b), 9(c) show the obtained clustering results. For comparative purposes, we also ran the OPTICSXi algorithm (which is a density based clustering technique) using these combined features. In Fig. 9(a), countries from the original AU's clustering were remapped into entirely new clusters. The West was made up of countries drawn from both Northern and Western African countries of AU's clustering. Similar reconfiguration were observed in the Central African cluster. These can be attributed to the influence of the geo-spatial features of the data set. To the South, clustering similar to that of the AU was observed; however, South Africa was singled out as a unit cluster. This is certainly due to the influence of data centre size feature, as South Africa has a significantly higher number of data centres than any other country in the Southern Africa region and the entire continent as a whole. Finally, the effect of population is obvious at the Eastern Cluster, as Ethiopia (with a significantly higher population) skews the clustering in its favour, resulting in the formation of a new cluster. The influences of the individual features on the overall clustering are also visible in the other results shown in Fig. 9(b) and 9(c). A comparative distribution of countries into clusters by K-Medoids and AU is shown in Fig. 10.

Fig. 10 shows the distribution of countries across the three clustering layouts, viz.: multi-feature clustering with k set to 5 (Fig. 9(a)), combined feature clustering with k set to 6 (Fig. 9(b)) and African Union clustering layout (Fig. 5(c)). In all three models, West Africa had the highest number of countries. Both K-Medoids models ($k = 5$ and 6), allocated 6 countries to East Africa, which is less than half of the AU's (at 13). With the $k = 5$ K-Medoids model, the North African cluster was completely removed, while $k = 6$ classified 5 countries as Northern countries versus AU's 8. K-Medoids grouped more countries into Central and Southern Africa than the AU did. A summary of the clustering algorithms and corresponding parameters is given on Table 4.

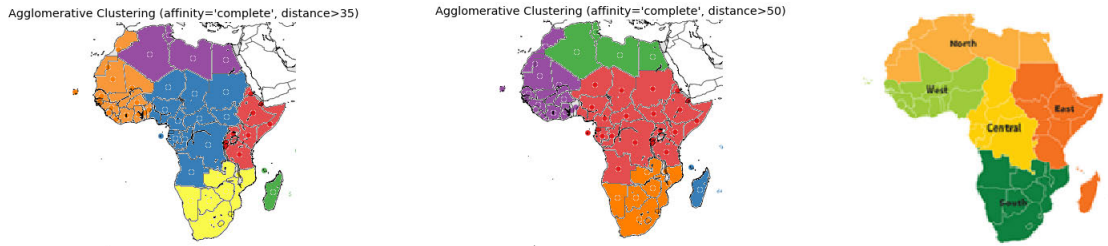


FIGURE 8. Hierarchical Agglomerative Clustering (HAC) – 35 and 50.

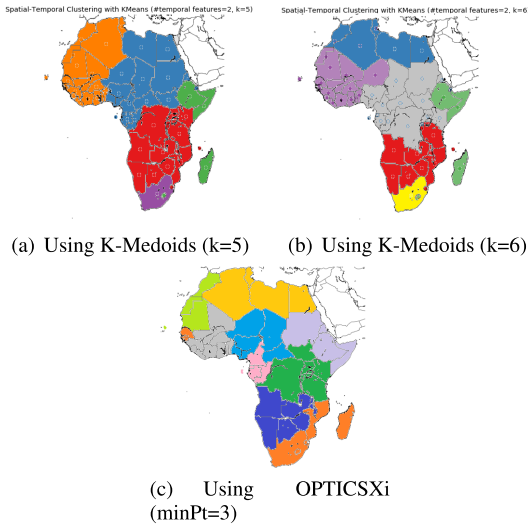


FIGURE 9. Combined feature clustering.

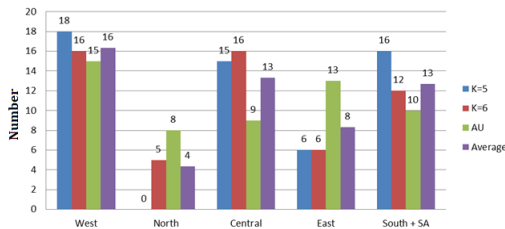


FIGURE 10. Comparison of country count per cluster.

B. ROUTING

Having put the countries into respective clusters, we now proceed with creating network(s) to interconnect countries across the continent. Our proposed model first creates an inner-network to link countries within clusters (intra-cluster routing); then creates a second high level network between clusters and to PCGs (inter-cluster routing). For comparison purposes, we also created a third network that treats each country as a cluster in itself and interconnects all unit countries to themselves. The Total Route Cost (TRC) is used as a metric to measure how good the selected route is. It is a composite metric derived from combining the population size, data centre count and geographical distance. Lower TRCs are desirable.

TABLE 4. Summary of clustering results.

Algorithm	Clustering Metric	Distance Metric	Clusters Count	Attributes
K-Means	$k = 5$	Euclidean	5	Lat. & Long.
K-Means	$k = 6$	Euclidean	6	Lat. & Long.
KMedoids	$k = 5$	Haversine	5	Lat. & Long.
KMedoids	$k = 6$	Haversine	6	Lat. & Long.
HAC	$Dist. = Complete Affinity$		6	Lat. & Long.
Multi Medoids + HAC	[K- $k = 5$ + HAC]	Haversine	5	Lat., Long., Pop. & DC count
Multi Medoids + HAC	[K- $k = 6$ + HAC]	Haversine	6	Lat., Long., Pop. & DC count
Multi [OPTICSXi]	$points = 3$	Haversine	10	Lat., Long., Pop. & DC count
Africa Union	NA	NA	5	Lat., Long. & Geo-political

C. INTRA-CLUSTER ROUTING

In this section, we present intra-cluster routing result for the various clustering layout obtained in section V-A. Obtained results give insights into the best network route from each African country to its closest continental gateway and how Africa as a continent can be interconnected. For each experiment 1000 ants were used. The cost metric used is a value representing the shortest path cost for a network which interconnects a country to its closest continental gateway (PCG) within each cluster. Less Desirable Trails (LDTs) are network paths which cut across the Sahara desert. Our model sought to discourage the use of these paths. For each country, unique network paths were obtained and are summarized on Tables 6–8.

1) NETWORK PATHS FOR $k = 5$ CLUSTERS

The combined feature clustering with $k = 5$, resulted in five clusters depicted in Fig. 9(a). Our proposed modified ACO-SP algorithm was used to obtain optimal paths between each country and the respective PCG for each of the clusters. The obtained simulation results are summarized on Table 6.

On the table, four clusters are presented instead of five, because we included South Africa in the Southern African cluster, rather than having it as a unit cluster. Since none of the PCGs, fell within the Western cluster, we set Ghana as the gateway. This was because Ghana (3) had the highest number

TABLE 5. Summary of routing costs.

	Multi Feature Clustering ($K = 5$)		Multi Feature Clustering ($K = 6$)		AU Clustering		Unit Clustering
	Country Count	Route Cost	Country Count	Route Cost	Country Count	Route Cost	
Western Cluster	18	43	16	27	15	27	NA
Northern Cluster	0	0	5	23	8	32	NA
Central Cluster	15	43	16	36	9	21	NA
Eastern Cluster	6	30	6	36	13	37	NA
Southern Cluster	16	37	12	31	10	21	NA
Intra-Cluster Total	55	153	55	152	55	138	55
Inter-Cluster Total	9 (PCGs)	61	10 (PCGs)	64	8 (PCGs)	54	8 (PCGs)
Continental Total (Intra + Inter)		214		216		191	131

of independent submarine landings of all the countries in the cluster (see Table 4). By doing this, we encouraged more traffic to flow through Ghana.

For the Central, Southern and Eastern African clusters, we set their PCGs to Cameroon, Djibouti, Nigeria; Kenya, South Africa; and Mauritius respectively. Similarly, we set Tunisia, Mali, Western Sahara and Mauritania as LDTs for the Western African cluster; Niger and Sudan as LDTs for the Central African cluster. No LDTs were set for the Southern and Eastern African cluster as there are no major deserts along the network path.

2) NETWORK PATHS FOR $k = 6$ CLUSTERS

The experiment was repeated for the second clustering layout shown in Fig. 9(b), where k was set to 6. Obtained results are summarized on Table 7. Similar to the previous results, South Africa was included in the Southern cluster, hence five clusters are shown instead of six. These are Western, Northern, Central, Southern and Eastern Africa clusters.

For the Western cluster, Ghana was again set as the gateway; while Mali, Western Sahara and Mauritania were set as LDTs. For the Northern African cluster, Algeria and Tunisia were set as the PCG and LDT respectively. For the Central African cluster, the PCGs were Cameroon, Djibouti, Kenya, Nigeria; while the Chad and Sudan were set as LDTs. For the Southern African cluster, South Africa was set as the PCG; while Mauritius was set as the PCG for the Eastern African cluster. For both the Southern and Eastern African clusters, no LDTs were set. The Table shows the optimal network path for each African country when $k = 6$ model was used.

3) NETWORK PATHS FOR AU CLUSTERS

For purpose of comparison, we also repeated the experiment to obtain the routing path for each country using the AU's clustering layout. Obtained results are as summarized on

Table 8. Five clusters namely Western, Northern, Central, Southern and Eastern Africa Clusters were presented.

Nigeria was set as the PCG for the Western cluster, while Mali and Niger were the LDTs. For the Northern African cluster, Algeria was the PCG used, while the LDTs consisted of Mauritania, Sudan, Tunisia, Western Sahara, being countries in/around the Sahara Desert. For the Central African cluster, Cameroon and Chad were set as the PCG and LDT respectively. For the Southern African cluster, South Africa was set as the PCG, while Kenya and Djibouti were the PCGs for the Eastern African cluster. For both the Southern and Eastern African Clusters, no LDTs were set. The Table shows the optimal network path for each country when Africa is split into clusters based on the AU's regional model.

D. COMPARING OPTIMAL PATHS

For each cluster, we also performed experiments to determine the "optimal path". For this work, an optimal path refers to a unique route, that traverses through all the countries within the cluster only once. This metric is useful in determining the best way to interconnect all countries within a cluster, while utilising the least amount of resources (e.g. Fibre optic or Ethernet cable). The result for each of the clusters within the three clustering scheme considered ($k = 5$, $k = 6$ and AU) are as follows:

- **$k = 5$: Western Africa Cluster.** We used the following experimental parameters: PCG: [Ghana]. LDTs: [Tunisia, Mali, Western Sahara and Mauritania]. Optimal network route: [Morocco, Algeria, Mauritania, Tunisia, Western Sahara, Cabo Verde, Senegal, Mali, Burkina Faso, Benin, Togo, Ghana, Gambia, Guinea Bissau, Liberia, Sierra Leone, Guinea, Cote d'Ivoire]. Total Route Cost (TRC): 43.
- **$k = 5$: Central Africa Cluster.** PCGs: [Cameroon, Djibouti, Nigeria]. LDTs: [Niger and Sudan]. Though Egypt and Libya are also located in the Sahara, they are not included in the LDT nor are they included as PCGs. This is because the influence of the stretch multiplier (for undesired paths) and boost multiplier (for desired paths) cancel each other out, hence no point including them in any of the lists. Optimal network route: [Republic of the Congo, Gabonese Republic, Equatorial Guinea, Sao Tome and Principe, Cameroon, Nigeria, Chad, Niger, Sudan, Libya, Egypt, Eritrea, Djibouti, South Sudan, Central African Republic]. Total Route Cost: 43.
- **$k = 5$: Southern Africa Cluster.** PCG: [Kenya, South Africa]. LDTs: [None]. We decide to include South Africa in this cluster and set it as one of the PCG, rather than have it as a unit cluster as shown in Fig. 9(a). Optimal Route: [Zimbabwe, Mozambique, Zambia, Malawi, Democratic Republic of Congo, Burundi, Rwanda, Uganda, Kenya, Tanzania, Comoros, South Africa, Eswatini, Botswana, Angola, Namibia]. Total Route Cost: 37.

- $k = 5$: **Eastern Africa Cluster**. PCG: [Mauritius]. LDT: [None]. In the absence of a PCG for this cluster, we chose Mauritius as it has a submarine landing and the highest number of data centres of all the other countries in the cluster. Optimal Route: [Lesotho, Madagascar, Mauritius, Seychelles, Somalia, Ethiopia]. Total Route Cost: 30.
- $k = 6$: **Western Africa Cluster**. PCG: [Ghana]. As before no country within the cluster met the requirement for PCG, hence we chose Ghana. LDTs: [Mali, Western Sahara and Mauritania]. Optimal Route: [Burkina Faso, Benin, Togo, Ghana, Cote d'Ivoire, Liberia, Sierra Leone, Guinea, Guinea Bissau, Gambia, Senegal, Cabo Verde, Western Sahara, Mauritania, Mali, Niger]. Total Route Cost = 27.33.
- $k = 6$: **Northern Africa Cluster**. PCG: [Algeria]. LDT: [Tunisia]. Egypt and Libya were not included for similar reasons as $k = 5$. Optimal Route: [Egypt, Libya, Algeria, Morocco, Tunisia]. Total Route Cost = 22.9.
- $k = 6$: **Central Africa Cluster**. PCGs: [Cameroon, Djibouti, Kenya, Nigeria]. LDTs: [Chad, Sudan]. Optimal Route: [Equatorial Guinea, Gabonese Republic, Republic of the Congo, Democratic Republic of Congo, Rwanda, Uganda, Kenya, South Sudan, Djibouti, Eritrea, Sudan, Chad, Central African Republic, Cameroon, Nigeria, Sao Tome and Principe]. Total Route Cost: 36.
- $k = 6$: **Southern Africa Cluster**. PCG: [South Africa]. As done earlier, South Africa is included in this cluster, to avoid having a unit cluster. LDT: [None]. Optimal Route: [Malawi, Mozambique, Zimbabwe, Eswatini, South Africa, Botswana, Namibia, Angola, Zambia, Burundi, Tanzania, Comoros]. Total Route Cost: 30.72.
- $k = 6$: **Eastern Africa Cluster**. PCG: [Mauritius]. LDT: [None]. Optimal Route: [Seychelles, Somalia, Ethiopia, Lesotho, Madagascar, Mauritius]. Total Route Cost: 36.
- **AU Western Africa Cluster**. PCG: [Nigeria]. LDTs: [Mali, Niger]. Optimal Route: [Senegal, Gambia, Guinea Bissau, Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Nigeria, Niger, Burkina Faso, Mali, Cabo Verde]. Total Route Cost: 26.7.
- **AU Northern Africa Cluster**. PCGs: [Algeria]. LDTs: [Mauritania, Sudan, Tunisia, Western Sahara]. Optimal Route: [Sudan, Egypt, Libya, Tunisia, Algeria, Morocco, Western Sahara, Mauritania]. Total Route Cost: 32.45.
- **AU Central Africa Cluster**. PCGs: [Cameroon]. LDTs: [Chad]. Optimal Route: [Central African Republic, Chad, Cameroon, Equatorial Guinea, Sao Tome and Principe, Gabonese Republic, Republic of the Congo, Democratic Republic of Congo, Burundi]. Total Route Cost: 21.
- **AU Southern Africa Cluster**. PCGs: [South Africa]. LDTs: [None]. Optimal Route: [Angola, Zambia, Malawi, Mozambique, Zimbabwe, Eswatini, Lesotho,

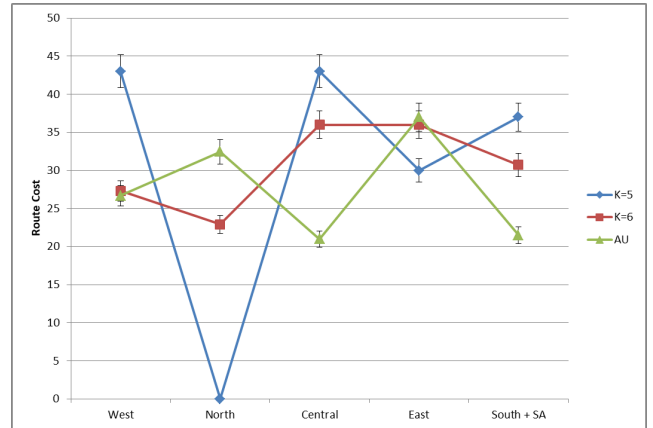


FIGURE 11. Intra-Cluster Route Cost.

South Africa, Botswana, Namibia]. Total Route Cost: 21.5.

- **AU Eastern Africa Cluster**. PCGs: [Kenya, Djibouti]. LDTs: [None]. Optimal Route: [South Sudan, Kenya, Uganda, Rwanda, Tanzania, Comoros, Madagascar, Mauritius, Seychelles, Somalia, Ethiopia, Djibouti, Eritrea]. Total Route Cost: 37.

Fig. 11 shows the total cost of traversing all countries within each of the clusters for the Multi-feature clustering (both $k = 5$ and $k = 6$) and the AU's clustering layout. The Northern cluster for $k = 5$ is zero because $k = 5$ did not create a Northern cluster – see Fig. 9(a). Comparatively, Fig. 11, shows that of all the three layouts, the AU's clustering layout resulted in the least traversal cost per cluster. Also routing across the $k = 5$ clusters generally cost more compared to $k = 6$, except in the East cluster where at 30, it was lower than $k = 6$'s 36 and AU's 37.

E. INTER-CLUSTER ROUTING

Having established a path between each cluster and its PCGs, there is the need to create a path across multiple clusters. This is the second (upper) level of our hierarchical clustering model. We again applied the modified ACO-SP algorithm to obtain the optimal route across the various clusters. Though this level represents routing across PCGs and all routes should ideally be desirable, we still included LDTs (which are paths that traverse the Sahara desert). Due to the small number of nodes, we reduced the number iteration to 50 and number of ants to 100. Obtained results are as follows:

- **Multi Feature Clustering with $k = 5$** PCGs: [Cameroon, Djibouti, Ghana, Kenya, Nigeria, Mauritius, South Africa] LDTs: [Egypt, Libya] Optimal Route: [Mauritius, South Africa, Cameroon, Nigeria, Ghana, Libya, Egypt, Djibouti, Kenya] Total Route Cost: 60.7
- **Multi Feature Clustering with $k = 6$** PCGs: [Cameroon, Djibouti, Ghana, Kenya, Nigeria, Mauritius, South Africa] LDTs: [Algeria, Egypt, Libya]

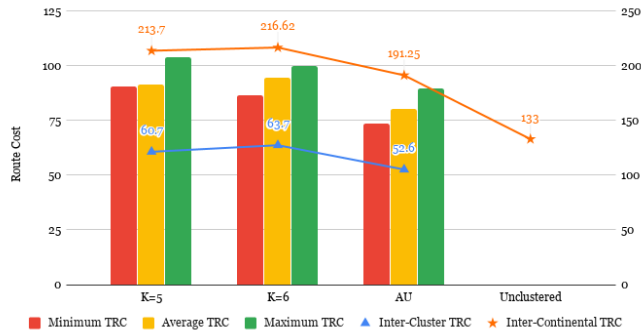


FIGURE 12. Inter-Cluster Traversal.

Optimal Route: [Kenya, Djibouti, Egypt, Libya, Algeria, Ghana, Nigeria, Cameroon, South Africa, Mauritius]
Total Route Cost: 63.7

- **AU Clustering** PCGs: [Cameroon, Djibouti, Kenya, Nigeria, South Africa,] LDTs: [Algeria, Egypt, Libya] Optimal Route: [Kenya, Djibouti, Egypt, Libya, Algeria, Nigeria, Cameroon, South Africa] Total Route Cost: 52.6

Fig. 12 shows the Total Routing Cost (TRC) for traversals across clusters. Of the three models, the AU’s clustering layout resulted in the lowest inter-cluster TRC at 52.6. This was followed by the multi feature clustering with k set to 5, while the option with k set to 6 cost the most to traverse the continent. Similar trends were observed for the average and maximum TRC. The minimum TRC for $k = 6$ was however lower than that of clustering with $k = 5$. This can be attributed to the fact that $k = 6$, created an additional cluster thereby, reducing the number of countries per cluster and by extension the traversal cost of some of its clusters.

F. CONTINENTAL ROUTING WITHOUT CLUSTERING

In this simulation we set each country as a unit entity rather a member of a cluster. Cameroon, Djibouti, Kenya, Nigeria and South Africa were set as PCGs; while least desired trail destinations included: Chad, Mali, Mauritania, Niger, Sudan, Tunisia and Western Sahara, as these are countries within the Sahara desert. Though Algeria, Egypt and Libya are also located in the Sahara, they are not included in the list of undesired nor are they included in the desired list. This is because the influence of the stench multiplier (for undesired paths) and boost multiplier (for desired paths) cancel each other out, hence no point including them in any of the lists. The simulation was run using 100 ants and obtained results are shown with the line graph in Fig. 12.

The line graph on Fig. 12 shows that the optimal route to traverse the entire continent would cost a TRC of 133. This was obtained by traversing the continent as follows: Nigeria, Cameroon, Equatorial Guinea, Gabonese Republic, Republic of the Congo, Sao Tome and Principe, Togo, Benin, Ghana, Cote d’Ivoire, Liberia, Sierra Leone, Guinea, Guinea Bissau, Gambia, Senegal, Cabo Verde, Mauritania, Western Sahara, Morocco, Algeria, Tunisia, Libya, Egypt, Sudan, Eritrea, Djibouti, Ethiopia, Somalia,

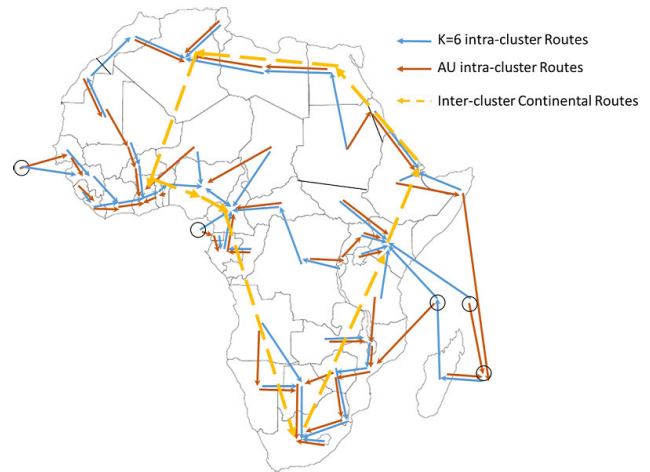


FIGURE 13. The African Continental Network Layout.

Kenya, Uganda, Rwanda, Burundi, Democratic Republic of Congo, Tanzania, Malawi, Mozambique, Zimbabwe, Zambia, Angola, Namibia, Botswana, South Africa, Lesotho, Eswatini, Madagascar, Mauritius, Comoros, Seychelles, South Sudan, Central African Republic, Chad, Niger, Mali, Burkina Faso.

When compared with the continental traversal obtained using the clustering based approaches, this unclustered approach resulted in the lowest TRC at 133. It was followed by the AU-based hierarchical clustering layout at 191.25, while the combined feature clustering with k set to 6 was the least effective approach requiring the highest TRC at 216.62. A summary of the routing costs is shown on Table 5 and depicted in Fig. 13.

VI. IMPLEMENTATION CONSIDERATIONS

In this section, we briefly discuss a potential use case of the model developed in this work; as well as some challenges that might hinder its implementation.

A. POTENTIAL USE CASE

The relevance or potential use case of this work, is in meeting the African Sustainable Development Goals (SDG) as defined in [50]. Of particular concern are those of: i.) good health and well being (through collaborative healthcare [1], [40]); ii.) industry, innovation and infrastructure (through scientific research and high performance computing such as Cloud computing); and iii.) sustainable cities and communities (smart and eco-cities [51]). These three goals are hinged on technologies of the fourth industrial revolution. An underlying network capable of supporting high bandwidth and data throughput is paramount to achieving them. Our Africa 3 model thus proposes a means of efficiently architecting such a network.

B. OPEN CHALLENGES

Thus far, deploying a pan African network for collaboration has remained elusive. A number of challenges are associated

TABLE 6. Routing path per country using K=5.

Source	Destination (PCGs)	Path	Cost
Benin	Algeria	Benin, Togo, Ghana, Burkina Faso, Mali, Mauritania, Western Sahara, Morocco, Algeria	20.77
Burkina Faso	Algeria	Burkina Faso, Mali, Mauritania, Western Sahara, Morocco, Algeria	17.10
Cabo Verde	Algeria	Cabo Verde, Gambia, Guinea Bissau, Senegal, Mauritania, Western Sahara, Morocco, Algeria	22.15
Cote d'Ivoire	Algeria	Cote d'Ivoire, Liberia, Sierra Leone, Guinea, Guinea Bissau, Gambia, Senegal, Mauritania, Western Sahara, Morocco, Algeria	22.44
Gambia	Algeria	Gambia, Senegal, Guinea Bissau, Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Burkina Faso, Mali, Mauritania, Western Sahara, Morocco, Algeria	28.07
Ghana	Algeria	Ghana, Togo, Benin, Burkina Faso, Mali, Mauritania, Western Sahara, Morocco, Algeria	21.38
Guinea	Algeria	Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Burkina Faso, Mali, Mauritania, Western Sahara, Morocco, Algeria	25.62
Guinea Bissau	Algeria	Guinea Bissau, Gambia, Senegal, Mauritania, Western Sahara, Morocco, Algeria	18.12
Liberia	Algeria	Liberia, Sierra Leone, Guinea, Guinea Bissau, Gambia, Senegal, Mauritania, Western Sahara, Morocco, Algeria	22.03
Mali	Algeria	Mali, Mauritania, Western Sahara, Morocco, Algeria	13.61
Mauritania	Algeria	Mauritania, Western Sahara, Morocco, Algeria	11.77
Morocco	Algeria	Morocco, Algeria	6.53
Senegal	Algeria	Senegal, Gambia, Guinea Bissau, Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Burkina Faso, Mali, Mauritania, Western Sahara, Morocco, Algeria	27.32
Sierra Leone	Algeria	Sierra Leone, Guinea, Guinea Bissau, Gambia, Senegal, Mauritania, Western Sahara, Morocco, Algeria	20.64
Togo	Algeria	Togo, Ghana, Benin, Burkina Faso, Mali, Mauritania, Western Sahara, Morocco, Algeria	21.72
Tunisia	Algeria	Tunisia, Algeria	5.90
Western Sahara	Algeria	Western Sahara, Mauritania, Senegal, Gambia, Guinea Bissau, Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Burkina Faso, Mali, Algeria	23.74
C.A.R	Cameroon	Central African Republic, Cameroon	4.41
Chad	Nigeria or Cameroon	Chad, Nigeria or Cameroon	7.71
Egypt	Djibouti	Egypt, Eritrea, Djibouti	13.09
Equatorial Guinea	Cameroon	Equatorial Guinea, Gabonese Republic, Republic of the Congo, Cameroon	5.94
Eritrea	Djibouti	Eritrea, Djibouti	3.33
Gabonese Republic	Cameroon	Gabonese Republic, Equatorial Guinea, Cameroon	4.79
Libya	Nigeria	Libya, Niger, Nigeria	14.39
Niger	Nigeria	Niger, Nigeria	6.17
Rep. of Congo	Cameroon	Republic of the Congo, Gabonese Republic, Equatorial Guinea, Cameroon	5.95
Sao Tome and Principe	Cameroon	Sao Tome and Principe, Gabonese Republic, Equatorial Guinea, Cameroon	6.52
South Sudan	Djibouti	South Sudan, Djibouti	7.61
Sudan	Cameroon	Sudan, Libya, Chad, Cameroon	21.84
Angola	South Africa	Angola, Namibia, Botswana, South Africa	14.71
Botswana	South Africa	Botswana, South Africa	5.37
Burundi	Kenya	Burundi, Rwanda, Uganda, Kenya	6.28
Comoros	Kenya	Comoros, Tanzania, Burundi, Rwanda, Uganda, Kenya	13.44
DRC	Kenya	Democratic Republic of Congo, Burundi, Rwanda, Uganda, Kenya	9.71
Eswatini	South Africa	Eswatini, South Africa	4.42
Malawi	South Africa	Malawi, Mozambique, Zimbabwe, Botswana, South Africa	14.69
Mozambique	South Africa	Mozambique, Malawi, Zambia, Zimbabwe, Eswatini, South Africa	15.81
Namibia	South Africa	Namibia, Botswana, South Africa	7.05
Rwanda	Kenya	Rwanda, Burundi, Kenya	5.72
Tanzania	Kenya	Tanzania, Rwanda, Burundi, Kenya	8.08
Uganda	Kenya	Uganda, Kenya	3.30
Zambia	Kenya	Zambia, DRC, Burundi, Rwanda, Burundi, Uganda, Kenya	15.31
Zimbabwe	South Africa	Zimbabwe, Eswatini, South Africa	9.03
Ethiopia	Mauritius	Ethiopia, Somalia, Mauritius	23.63
Lesotho	Mauritius	Lesotho, Mauritius	17.88
Madagascar	Mauritius	Madagascar, Mauritius	6.31
Seychelles	Mauritius	Seychelles, Mauritius	11.39
Somalia	Mauritius	Somalia, Ethiopia, Mauritius	23.63

with implementing such a network, some of which are as follows:

- **Policies:** The policies (or lack of) are a major deterrent to collective growth of Africa. A number of policies on

TABLE 7. Routing path per country using K=6.

Source	Destination (PCGs)	Path	Cost
Benin	Ghana	Benin, Togo, Ghana	2.46
Burkina Faso	Ghana	Burkina Faso, Ghana	3.21
Cabo Verde	Ghana	Cabo Verde, Senegal, Gambia, Guinea Bissau, Guinea, Liberia, Cote d'Ivoire, Ghana	15.94
Cote d'Ivoire	Ghana	Cote d'Ivoire, Ghana	2.56
Gambia	Ghana	Gambia, Senegal, Guinea Bissau, Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana	11.22
Guinea	Ghana	Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana	7.10
Guinea Bissau	Ghana	Guinea Bissau, Gambia, Senegal, Sierra Leone, Liberia, Cote d'Ivoire, Ghana	10.79
Liberia	Ghana	Liberia, Cote d'Ivoire, Ghana	4.81
Mali	Ghana	Mali, Burkina Faso, Ghana	6.99
Mauritania	Ghana	Mauritania, Mali, Burkina Faso, Ghana	10.84
Niger	Ghana	Niger, Ghana	9.12
Senegal	Ghana	Senegal, Gambia, Guinea Bissau, Guinea, Sierra Leone, Cote d'Ivoire, Ghana	10.44
Sierra Leone	Ghana	Sierra Leone, Guinea, Liberia, Cote d'Ivoire, Ghana	7.68
Togo	Ghana	Togo, Ghana	1.37
Western Sahara	Ghana	Western Sahara, Mauritania, Mali, Burkina Faso, Ghana	13.86
Egypt	Algeria	Egypt, Libya, Algeria	15.72
Libya	Algeria	Libya, Algeria	8.75
Morocco	Algeria	Morocco, Algeria	6.53
Tunisia	Algeria	Tunisia, Algeria	5.90
Chad	Nigeria	Chad, Nigeria	7.71
C.A.R	Cameroon	Central African Republic, Cameroon	4.41
DRC	Kenya	DRC, Rwanda, Uganda, Kenya	9.23
Equatorial Guinea	Cameroon	Equatorial Guinea, Gabonese Republic, Republic of the Congo, Cameroon	5.94
Eritrea	Djibouti	Eritrea, Djibouti	3.33
Gabonese Republic	Cameroon	Gabonese Republic, Equatorial Guinea, Cameroon	4.79
Rep. of Congo	Cameroon	Republic of the Congo, Gabonese Republic, Equatorial Guinea, Cameroon	5.95
Rwanda	Kenya	Rwanda, Kenya	4.97
Sao Tome and Principe	Cameroon	Sao Tome and Principe, Gabonese Republic, Equatorial Guinea, Cameroon	6.52
South Sudan	Kenya	South Sudan, Kenya	6.52
Sudan	Djibouti	Sudan, Eritrea, Djibouti	8.09
Uganda	Kenya	Uganda, Kenya	3.30
Angola	South Africa	Angola, Namibia, Botswana, South Africa	14.71
Botswana	South Africa	Botswana, South Africa	5.37
Burundi	South Africa	Burundi, Tanzania, Malawi, Mozambique, Zimbabwe, Botswana, South Africa or Burundi, Tanzania, Malawi, Mozambique, Zimbabwe, Eswatini, South Africa	22.0
Comoros	South Africa	Comoros, Mozambique, Malawi, Zambia, Eswatini, South Africa or Comoros, Malawi, Mozambique, Zimbabwe, Botswana, Namibia, South Africa	21.68
Eswatini	South Africa	Eswatini, South Africa	4.41
Malawi	South Africa	Malawi, Mozambique, Zimbabwe, Botswana, South Africa	14.47
Mozambique	South Africa	Mozambique, Zimbabwe, Botswana, South Africa	11.86
Namibia	South Africa	Namibia, Botswana, South Africa	8.05
Tanzania	South Africa	Tanzania, Burundi, Zambia, Zimbabwe, Botswana, South Africa	21.17
Zambia	South Africa	Zambia, Malawi, Mozambique, Zimbabwe, Eswatini, South Africa	15.95
Zimbabwe	South Africa	Zimbabwe, Eswatini, South Africa	9.03
Ethiopia	Mauritius	Ethiopia, Somalia, Mauritius	23.63
Lesotho	Mauritius	Lesotho, Mauritius	17.88
Madagascar	Mauritius	Madagascar, Mauritius	6.31
Seychelles	Mauritius	Seychelles, Mauritius	11.39
Somalia	Mauritius	Somalia, Ethiopia, Mauritius	23.63

Agriculture and food security, health and education are in place but have remained dormant or ill-implemented. From a data and ICT standpoint, the Protection of Personal Information Act and data sovereignty [47]–[49] are fundamental issues that need to be addressed. They are concerned with data legislation on where, who and how data should be stored or accessed across national boundaries. These need to be clearly agreed upon before a truly collaborative data network, such as Africa 3 can be implemented.

- **Technical Expertise:** Beyond policies, technical experts would be needed to design, deploy and operate a continental network such as Africa 3. Unfortunately, as of today, there is still a wide gap in technical know-how between Africa and the developed world. This is especially true for ICT and related fields. In order to achieve a sustainable continental network, Africans need to be up-skilled in emerging fields such areas next generation networks, Cloud computing and data engineering. There is also the need to focus on Cyber

TABLE 8. Routing path per country using AU's model.

Source	Destination (PCGs)	Path	Cost
Benin	Nigeria	Benin, Nigeria	2.36
Burkina Faso	Nigeria	Burkina Faso, Ghana, Togo, Benin, Nigeria	7.90
Cabo Verde	Nigeria	Cabo Verde, Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Nigeria	20.96
Cote d'Ivoire	Nigeria	Cote d'Ivoire, Ghana, Togo, Benin, Burkina Faso, Nigeria	11.36
Gambia	Nigeria	Gambia, Senegal, Guinea Bissau, Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Nigeria or Gambia, Senegal, Guinea Bissau, Sierra Leone, Guinea, Cote d'Ivoire, Ghana, Togo, Benin, Nigeria	16.77
Ghana	Nigeria	Ghana, Togo, Benin, Nigeria	6.07
Guinea	Nigeria	Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Nigeria	12.89
Guinea Bissau	Nigeria	Guinea Bissau, Gambia, Senegal, Guinea, Sierra Leone, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Nigeria	16.76
Liberia	Nigeria	Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Nigeria	10.80
Mali	Nigeria	Mali, Burkina Faso, Ghana, Togo, Benin, Nigeria	11.18
Niger	Nigeria	Niger, Nigeria	6.17
Senegal	Nigeria	Senegal, Gambia, Guinea Bissau, Guinea, Sierra Leone, Liberia, Ghana, Togo, Benin, Nigeria	16.33
Sierra Leone	Nigeria	Sierra Leone, Guinea, Liberia, Cote d'Ivoire, Ghana, Togo, Benin, Nigeria	13.61
Togo	Nigeria	Togo, Benin, Nigeria	4.71
Burundi	Cameroon	Burundi, Democratic Republic of Congo, Cameroon	12.08
C.A.R	Cameroon	Central African Republic, Cameroon	4.41
Chad	Cameroon	Chad, Cameroon	7.71
DRC	Cameroon	Democratic Republic of Congo, Burundi, Central African Republic, Cameroon	12.91
Equatorial Guinea	Cameroon	Equatorial Guinea, Gabonese Republic, Republic of the Congo, Cameroon	5.94
Gabonese Republic	Cameroon	Gabonese Republic, Equatorial Guinea, Sao Tome and Principe, Cameroon	6.85
Rep. of Congo	Cameroon	Republic of the Congo, Gabonese Republic, Equatorial Guinea, Cameroon	5.95
Rwanda	Kenya	Rwanda, Kenya	4.97
Sao Tome and Principe	Cameroon	Sao Tome and Principe, Gabonese Republic, Equatorial Guinea, Cameroon	6.52
Angola	South Africa	Angola, Botswana, Namibia, South Africa	15.97
Botswana	South Africa	Botswana, South Africa	5.37
Eswatini	South Africa	Eswatini, South Africa	4.41
Lesotho	South Africa	Lesotho, South Africa	2.24
Malawi	South Africa	Malawi, Mozambique, Zimbabwe Botswana, South Africa	14.47
Mozambique	South Africa	Mozambique, Zimbabwe, Botswana, South Africa	11.86
Namibia	South Africa	Namibia, Botswana, South Africa	8.05
Zambia	South Africa	Zambia, Malawi, Mozambique, Zimbabwe, Eswatini, South Africa	15.95
Zimbabwe	South Africa	Zimbabwe, Eswatini, South Africa	9.03
Comoros	Kenya	Comoros, Kenya	9.70
Ethiopia	Djibouti	Ethiopia, Djibouti	2.97
Eritrea	Djibouti	Eritrea, Djibouti	3.33
Madagascar	Kenya	Madagascar, Comoros, Tanzania, Kenya	16.51
Mauritius	Kenya	Mauritius, Madagascar, Comoros, Tanzania, Kenya	21.44
Rwanda	Kenya	Rwanda, Uganda, Kenya	5.53
Seychelles	Kenya	Seychelles, Kenya	10.76
Somalia	Djibouti or Kenya	Somalia, Djibouti or Somalia, Kenya	5.4 or 5.5
South Sudan	Kenya	South Sudan, Kenya	6.58
Tanzania	Kenya	Tanzania, Kenya	5.47
Uganda	Kenya	Uganda, Kenya	3.30

security and threats, as attacks on a continental network could have far reaching consequences.

- **Electricity:** Without a doubt, electricity is a fundamental requirement for any data network. Though at a continental level, one would assume that electricity would be available to power Africa 3; the actual problem lies at the edges or peripherals of this network. Here, the term edges refer to the individual countries connected to Africa 3, and which serve as “data prosumers”. As of today, electricity remains one of the biggest bottlenecks to sustainable development in many edges. Hence, for a continental network

such as Africa 3 to be truly useful, African countries would require widespread energy reforms. This includes exploring alternate, renewable and sustainable sources of electricity such as wind, solar, wave and hydro.

- **Poverty:** Though Africa 3 is being proposed to help accelerate development across Africa, the model in itself has to be viable / self-sustainable. Users therefore have to pay to use it. Sadly, over 60% of the sub-Saharan populace still live in poverty [52]. Many families are unable to afford good food, healthcare or education for their children. Poverty translates to little purchasing

power, weak economies and ultimately little or no funds to sustain the continental network. If deployed as a non-commercial entity and left unfunded, such continental network would collapse / die-off at its infancy.

VII. CONCLUSION

Reports of the United Nations on global human development index, shows that more than 75% of countries in Africa are underdeveloped. These countries lack vital infrastructure for health care, education, transportation, electricity etc. This situation is further exacerbated by the high level of poverty in many of these countries. Collaboration enabled by Information and communication Technology (ICT) could prove to be a vital tool in fast-tracking development across Africa. In achieving this, an underlying robust intra-continental network structure must first be put in place; upon which concepts such as tele-medicine, remote learning, e-governance, smart cities and federated Cloud processing and storage can be built. The intricacies of such a network are often enormous as factors such as geo-location, population size, socio-economic and politics need to be considered. These therefore complicate the process. In this work, we proposed a network model for inter-connecting African countries. This two-layered hierarchical model takes into consideration the population size, the geographical locations, the number of independent sub-marine cable landings and the number of high performance computing infrastructure available in each African country into consideration. At the first layer, African countries are grouped into functional clusters. Four clustering algorithms were considered, viz. K-Means, K-Medoids, Hierarchical clustering and OPTICS-Xi. These were bench-marked against the Africa Union's clustering (AU) and obtained results show that the K-Medoids gave results closest to the AU. Upon successfully clustering the countries, a modified variant of the Ant Colony Optimization algorithm called Ant Colony Optimization with Stench Pheromone was used to obtain least cost network paths within and between the clustered. Stench Pheromones were used to dissuade paths that traversed the Sahara desert. Simulation results show that indeed an optimal network path can be obtained that traversed all African nations, and that such a network would be most efficiently designed using the Africa Union's clustering layout. Implementation challenges of this network model in Africa were then discussed.

Though a robust network model for Africa was proposed in this work; the authors have assumed that perfect line of sight exists when linking countries. Furthermore, ground elevation, weather conditions and geo-political constraints such as language diversity and trans-border rifts were not considered. Finally, though a network has been proposed, the exact technology in terms of transmission media and redundancy were also not considered. These are potential areas that might be considered for future works.

ACKNOWLEDGMENT

The authors would like to thank Prof. Biswanath Mukerjee and Dr. Clement Nyirenda for their reviews and comments, both of which helped improve this work.

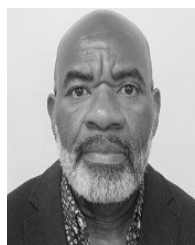
REFERENCES

- [1] O. O. Ajayi, A. B. Bagula, and K. Ma, "Fourth industrial revolution for development: The relevance of cloud federation in healthcare support," *IEEE Access*, vol. 7, pp. 185322–185337, 2019.
- [2] *DataCentreMap*. Accessed: Mar. 21, 2020. [Online]. Available: <https://www.datacentermap.com>
- [3] A. Huang, S. Sarkar, and B. Mukherjee, "Africa two: A proposal for a concentric two-ring network for the African continent," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 6, pp. 2–11, Aug. 2008.
- [4] Worldometers. (2019). *African Countries by Population*. Accessed: Nov. 9, 2019. [Online]. Available: <https://www.worldometers.info/population/countries-in-africa-by-population/>
- [5] GADM Maps and Data. Accessed: Mar. [Online]. Available: <https://www.gadm.org/index.html>
- [6] C. DiMinico and J. Jew, "Telecommunications infrastructure standard for data centers ANSI/TIA-942," Tech. Rep., 2016.
- [7] African Union. (2019). *Member States*. Accessed: Nov. 10, 2019. [Online]. Available: <https://www.au.int/en/member-states/countryprofiles2>
- [8] West Africa Brief. *The Six Regions of the African Union*. Accessed: Feb. 26, 2020. [Online]. Available: <http://www.west-africa-brief.org/content/en/six-regions-african-union>. Accessed 26/02/2020.
- [9] P. Apparicio, M. Riva, and A. Seguin, "A comparison of two methods for clustering trajectories: A case study on neighborhood poverty at the intra-metropolitan level in Montreal," *Cybergeo: Eur. J., Geography, Espace, Society, Territory*, document 727, 2015, doi: 10.4000/cybergeo.27035.
- [10] M. Ester, H. J. Kriegel Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowl. Discovery Data Mining (KDD)*, 1996, pp. 226–231.
- [11] F. Murtagh and P. Contreras, "Methods of hierarchical clustering," 2011, *arXiv:1105.0121*. [Online]. Available: <http://arxiv.org/abs/1105.0121>
- [12] *Seacom Network*. Accessed: Mar. 11, 2020. [Online]. Available: <http://www.seacom.com/network>
- [13] D. Xu and Y. Tian, "A comprehensive survey of clustering algorithms," *Ann. Data Sci.*, vol. 2, no. 2, pp. 165–193, 2015.
- [14] D. Cheng, Q. Zhu, J. Huang, Q. Wu, and L. Yang, "A hierarchical clustering algorithm based on noise removal," *Int. J. Mach. Learn. Cybern.*, vol. 10, no. 7, pp. 1591–1602, Jul. 2019.
- [15] Y. Chen, S. Tang, S. Pei, C. Wang, J. Du, and N. Xiong, "DHeat: A density heat-based algorithm for clustering with effective radius," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 48, no. 4, pp. 649–660, Apr. 2018.
- [16] J. Jiang, Y. Chen, X. Meng, L. Wang, and K. Li, "A novel density peaks clustering algorithm based on k nearest neighbors for improving assignment process," *Phys. A: Stat. Mech. Appl.*, vol. 523, pp. 702–713, Jun. 2019.
- [17] Y. Chen, S. Tang, N. Bouguila, C. Wang, J. Du, and H. Li, "A fast clustering algorithm based on pruning unnecessary distance computations in DBSCAN for high-dimensional data," *Pattern Recognit.*, vol. 83, pp. 375–387, Nov. 2018.
- [18] *Africa Network Connections–SAT3/WASC/SAFE*. [Online]. Available: <https://www.cmcnetworks.net/sat3-wasc-safe.html>. Accessed 11/03/20.
- [19] *Submarine Cables*. Accessed: Mar. 11, 2020. [Online]. Available: <https://www.submarinecablemap.com>
- [20] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization," *IEEE Comput. Intell. Mag.*, col. 1, no. 4, pp. 28–39, Nov. 2006.
- [21] N. Kaur, U. Kaur, and D. Singh, "K-Medoid clustering algorithm-a review," *Int. J. Comput. Appl. Technol.*, vol. 1, no. 1, p. 2349, 2014.
- [22] L. Morissette and S. Chartier, "The k-means clustering technique: General considerations and implementation in mathematica," *Tuts. Quant. Methods Psychol.*, vol. 9, no. 1, pp. 15–24, Feb. 2013.
- [23] I. Dokmanic, R. Parhizkar, J. Ranieri, and M. Vetterli, "Euclidean distance matrices: Essential theory, algorithms, and applications," *IEEE Signal Process. Mag.*, vol. 32, no. 6, pp. 12–30, Nov. 2015.
- [24] C. C. Robusto, "The cosine-haversine formula," *Amer. Math. Monthly*, vol. 64, no. 1, pp. 38–40, 1957.

- [25] M. Ankerst, M. Breunig, H. Kriegel, and J. Sander, "OPTICS: ordering points to identify the clustering structure," *ACM SIGMOD Rec.*, vol. 28, no. 2, pp. 49–60, 1999, doi: [10.1145/304181.304187](https://doi.org/10.1145/304181.304187).
- [26] Optics. Accessed: Mar. 10, 2020. [Online]. Available: <http://www.rdr.io/cran/dbscan/man/optics.html>
- [27] S. Patro and K. Sahu, "Normalization: A preprocessing stage," Department Comp. Sci. Eng. Intell. Transp. (CES & IT), Veer SurendraSai Univ. Technol. (VSSUT), Odisha, India, 2015.
- [28] I. B. Mohamad and D. Usman, "Standardization and its effects on K-means clustering algorithm," *Res. J. Appl. Sci., Eng. Technol.*, vol. 6, no. 17, pp. 3299–3303, 2013, doi: [10.19026/rjaset.6.3638](https://doi.org/10.19026/rjaset.6.3638).
- [29] E. Schubert and A. Zimek, "ELKI: A large open-source library for data analysis—ELKI release 0.7.5 'Heidelberg'," 2019, *arXiv:1902.03616*. [Online]. Available: <http://arxiv.org/abs/1902.03616>
- [30] F. Pedregosa, G. Varoquaux, and A. Gramfort, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Oct. 2011.
- [31] Scikit Yellowbrick, *Elbow Method*. Accessed: Mar. 18, 2020. [Online]. Available: <http://www.scikit-yb.org/en/latest/api/cluster/elbow.html>
- [32] F. Di Caro, F. Ducatelle, and L. M. Gambardella, "AntHocNet: an adaptive nature-inspired algorithm for routing in mobile ad hoc networks," *Special Issue Self-Organisation Mobile Netw.*, vol. 16, no. 5, pp. 443–455, Sep./Oct. 2005.
- [33] G. Caro and M. Dorigo, "Ant colonies for adaptive routing in packet-switched communications networks," in *Proc. ACM Intl. Conf. Parallel Problem Solving Nature*, 1998, pp. 673–682.
- [34] S. Kamali and J. Opatrny, "A position based ant colony routing algorithm for mobile ad-hoc networks," *J. Netw.*, vol. 3, no. 4, pp. 31–41, Apr. 2008.
- [35] Z. Cong, B. De Schutter, and R. Babuška, "Ant colony routing algorithm for freeway networks," *Transp. Res. C, Emerg. Technol.*, vol. 37, pp. 1–19, Dec. 2013, doi: [10.1016/j.trc.2013.09.008](https://doi.org/10.1016/j.trc.2013.09.008).
- [36] Z. Cong, B. De Schutter, and R. Babuska, "A new ant colony routing approach with a trade-off between system and user optimum," in *Proc. 14th Int. IEEE Conf. Intell. Transp. Syst. (ITSC)*, Oct. 2011, pp. 1369–1374.
- [37] J. E. Bell and P. R. McMullen, "Ant colony optimization techniques for the vehicle routing problem," *Adv. Eng. Informat.*, vol. 18, no. 1, pp. 41–48, Jan. 2004, doi: [10.1016/j.aei.2004.07.001](https://doi.org/10.1016/j.aei.2004.07.001).
- [38] United Nations Development Program, *Human Development Indices and Indicators. Human Development Report 2019*. HDRO (Human Development Report Office) United Nations Development Programme, New York, NY, USA, 2019, pp. 22–25.
- [39] M. Foley, *The Role and Status of National Research and Education Networks in Africa*. 2016.
- [40] M. Mandava, C. Lubamba, A. Ismail, H. Bagula, and A. Bagula, "Cyber-healthcare for public healthcare in the developing world," in *Proc. IEEE Symp. Comput. Commun. (ISCC)*, Messina Italy, Jun. 2016, pp. 14–19.
- [41] J. Adams, "Collaborations: The rise of research networks," *Nature*, vol. 490, no. 7420, pp. 335–336, Oct. 2012.
- [42] W. C. Marra and J. Schesser, "Africa ONE: The africa optical network," *IEEE Commun. Mag.*, vol. 34, no. 2, pp. 50–57, 1996.
- [43] P. Mell and T. Grance, "The NIST Definition of Cloud Computing," NIST Special Publication 800-145, Gaithersburg, MD, USA, Whitepaper, 2011, p. 7.
- [44] J. Djamen, D. Ramazani, and S. Some (1995), "Networking in Africa: An unavoidable evolution towards the Internet," Dept. d'Informatique et de Recherche Operationnelle (IRO), Universite de Montreal, Montreal, QC, Canada, Tech. Rep. 937, 1995.
- [45] C. Rey-Moreno. (2017). *Supporting the creation and scalability of affordable access solutions: Understanding Community Networks in Africa*. Accessed: Apr. 1, 2020. [Online]. Available: <https://internetsociety.org/wp-content/.../May2017-1.pdf>
- [46] J. Chavula, N. Feamster, A. Bagula, and H. Suleman, "Quantifying the effects of circuitous routes on the latency of intra-Africa Internet traffic: A study of research and education networks," in *Proc. Int. Conf. e-Infrastruct. e-Services Developing Countries*, 2014, pp. 64–73.
- [47] M. De Bruyn, "The protection of personal information (POPI) act: Impact on South Africa," *Int. Business Econ. Res. J.*, vol. 13, no. 6, 1315, 2014.
- [48] P. Sello, A. Bagula, and O. Ajayi, "Laws and Regulations on Big Data Management: The Case of South Africa," in *Intl. Conf. on e-Infrastructure and e-Services for Developing Countries*. Cham, Switzerland: Springer, 2019, pp. 169–179.
- [49] N. Kushwaha, P. Roguski, and B. W. Watson, "Up in the air: Ensuring government data sovereignty in the cloud," in *Proc. 12th Int. Conf. Cyber Conflict (CyCon)*, May 2020, pp. 43–61.
- [50] L. Henao, L. Moyer, and P. Namakula, *Africa 2030: How Africa Can Achieve the Sustainable Development Goals*. Kigali, Rwanda: Sustainable Development Goal Centre for Africa (SDGC/A), 2017.
- [51] S. Kyakulumbye, S. Pather, and A. Bagula, "Smart city Citizens' service provision using participatory design and participatory sensing: Lessons for developing cities," in *Proc. IST-Africa Conf. (IST-Africa)*, May 2020, pp. 1–12.
- [52] WHO. (2014). *The health of the people: what works. The African Regional Health Report 2014*. Accessed: Oct. 5, 2020. [Online]. Available: <https://www.apps.who.int/iris/bitstream/10665/137377/4/9789290232612.pdf>



books, and proceedings of international conferences.



Antoine B. Bagula (Member, IEEE) received the dual M.Sc. degrees in computer engineering and computer science from the Université catholique de Louvain (UCL), Belgium, and Stellenbosch University (SUN), South Africa, respectively, and the Ph.D. degree in communication systems from the KTH Royal Institute of Technology, Sweden. He is currently a Full Professor and the Head of the Department of Computer Science, University of the Western Cape, Cape Town, South Africa, where he also leads the Intelligent Systems and Advanced Telecommunication (ISAT) Laboratory. He is a well published Scientist in his research field. His research interests include the Internet of Things, cloud/fog computing, and next generation networks, including 4G/5G. For the last ten years, he has applied a combination of these research interests to the design and implementation of cyber physical systems frameworks for health, energy, water, and pollution.



Hloniphani C. Maluleke received the B.Sc. degree in computer science, in 2015. His M.Sc. thesis was focused on providing 5G communication in rural and under-served communities of South Africa. He is currently pursuing the Ph.D. degree with the University of the Western Cape, Cape Town, South Africa, with a research focus on application of cyber-physical systems in pollution monitoring, under the supervision of Prof. B. Bagula. His research interests include 5G communications, machine learning, CPS, and 4IR for rural and under-served areas of South Africa.

...